JOURNAL OF COMPLEXITY 6, 41–66 (1990)

s-Numbers in Information-Based Complexity

P. MATHÉ

Akademie der Wissenschaften der DDR, Karl-Weierstrass-Institut für Mathematik, Mohrenstraße 39, Berlin, DDR-1086

Received May 16, 1988

We shall study maximal errors of approximating linear problems. As possible classes of information operators the classes of arbitrary, continuous (nonlinear), or continuous linear information operators are considered. Algorithms also may be arbitrary, continuous (nonlinear), or linear. We focus our interest on two natural questions: (a) For what problems (the dependence on the underlying Banach spaces turns out to be crucial) do different classes of information or algorithms, respectively, yield the same quality of approximation? (b) What are the maximal differences in the errors of different classes? Both questions are treated in both the worst-case and average-case settings. Therefore the paper is divided into Parts A and B. For the study of the worst-case setting the notion of s-scales turns out to be powerful. An appropriate approach is also suggested for the average-case setting. Using the ideas of s-scales and function analytic methods we reprove some known results and obtain some new ones, thus answering questions posed in several papers on this subject. © 1990 Academic Press, Inc.

INTRODUCTION

The aim of this paper is to give a detailed analysis of various (approximate) methods for solving linear problems in the sense of Information-Based Complexity. The results can be of interest in at least two aspects. For those working in functional analysis we give interpretations of results of the theory of s-numbers in (abstract) numerical analysis, as outlined below. For those working in numerical analysis we give applications of function analytic methods to answer problems arising very naturally in the theory of optimal algorithms.

Toward this purpose we shall follow a common scheme. Fix for a moment Banach spaces E, F and a linear operator S acting between E and F. Given x in the unit ball B_E of E we approximate Sx in the following way.

0885-064X/90 $3.00

First we choose information $N(x)$, which is to lie in some finite-dimensional space, and then, using $N(x)$, we apply an algorithm φ to obtain $\varphi(N(x)) \in F$ as an approximation for Sx. Generally, the result will not be exact; thus we have an (individual) error $Sx - \varphi(N(x))$ at x. At this point there are several possibilities for measuring the quality of N and φ.

Part A is devoted to the study of the worst-case error $e(S, N, \varphi)$. By this we mean that we take the $\sup\{\|Sx - \varphi(N(x))\| : x \in B_E\}$ for fixed N and φ. Now we are able to compare the quality of different algorithms using the same information or even the quality of different information operators. More precisely, we shall do this for N and φ varying in some natural classes, i.e., arbitrary, continuous, and continuous linear mappings.

Part B is devoted to the study of the average-case error. Here we choose a symmetric Radon measure μ on E. The average-case error $e(S, N, \varphi, \mu)$ is then defined to be $\int \|Sx - \varphi(N(x))\| d\mu(x)$, again for fixed N and φ. Proceeding as in Part A, we will be able to compare different information operators and algorithms.

Both cases (and some others) have been subject to investigations during the last years. We shall not give complete references, but restrict ourselves to "A General Theory of Optimal Algorithms" (Traub and Woźniakowski, 1980) and "Information-Based Complexity" (Traub, Wasilkowski, and Woźniakowski, 1988) and the survey papers "Information-based complexity" (Woźniakowski, 1986) and "Recent developments in information-based complexity" (Packel and Woźniakowski, 1987). Comprehensive bibliographies can be found in each of these references. This paper was initiated by a recent result of Kacewicz and Wasilkowski (1986) in "How powerful is continuous nonlinear information?" They showed, for operators acting between Hilbert spaces, that for every continuous (nonlinear) information there is linear information with the same error in the worst-case setting, while there is a big difference in the average-case setting. We shall derive these results from a more general approach. Namely, in Theorem 1 and Corollary 2 we establish a relation to the theory of s-numbers which have been also studied extensively. We mention the monographs "Operator Ideals" and "Eigenvalues and s-Numbers" (Pietsch, 1978, 1987). Once we have seen this, the result of Kacewicz and Wasilkowski is a consequence of a general result on s-numbers. Moreover, we are now in a position to prove and reprove several other results. We pay attention to two questions. First, in what situations can different classes yield the same errors? Second, what are the maximal differences in the errors of different classes? Both questions are made precise in Sections 2 and 3 of Part A. For the average-case setting, problems like those stated above are less studied. We outline our approach in Part B. To have the right tools, we introduce s-number-like mappings in Section 4. A first approach like this has been proposed by Micchelli (1984) in "Orthogonal projections are optimal algorithms." Interestingly enough, there are big differences in the behavior of the worst-case and average-case errors. The relations are studied in detail in Sections 5 and 6. Open questions are also mentioned.

I wish to express my gratitude to S. Heinrich, Berlin, who not only suggested the topic but also paid instant attention to it and was a source of fruitful ideas.

PART A: THE WORST-CASE SETTING

1. Notation and General Results

In this part the notation is standard or adopted from Traub and Woźniakowski (1980) (abbreviated as GTOA) or Pietsch (1978, 1987).

Given $m \in \mathbb{N}$ define $l_p^m := [\mathbb{R}^m, \|\cdot\|_p]$, with

$$\|(x_1, x_2, \ldots, x_m)\| := \begin{cases} (\sum |x_i|^p)^{1/p}, & 1 \le p < \infty \\ \max\{|x_i|, 1 \le i \le m\}, & p = \infty. \end{cases}$$

Given a compact Hausdorff space K and a Banach space F, we denote by $C(K, F)$ the space of all F-valued continuous functions on K with its usual norm. For a given measure space $[\Omega, \mathcal{F}, \mu]$, Banach space F, and fixed $1 \le p \le \infty$ we denote by $L_p(\Omega, \mathcal{F}, \mu, F)$, $L_p(\mu, F)$ if $[\Omega, \mathcal{F}]$ is given $L_p(\mu)$ if $F = \mathbb{R}$, the closure of the space of all equivalence classes of F-valued step functions under the norm

$$\|f\|_p := \begin{cases} (\int \|f(\omega)\|_F^p d\mu(\omega))^{1/p}, & 1 \le p < \infty \\ \text{ess-sup}\{\|f(\omega)\|_F, \omega \in \Omega\}, & p = \infty. \end{cases}$$

Given Banach spaces E, F we denote by $L(E, F)$ the linear space of all bounded linear operators from E to F, equipped with the operator norm. For a fixed operator $T \in L(E, F)$ let $\ker T := \{x \in E : Tx = 0\}$ and $\operatorname{Im} T := \{y \in F : \exists x \in E \text{ and } Tx = y\}$ denote the kernel and image of T, respectively.

We are interested in special operators. An operator T is of finite rank if $\dim \operatorname{Im} T < \infty$. An operator $P \in L(E, E)$ is called a projection if $P^2 = P$. An operator $J \in L(E, F)$ is said to be a metric injection if $\|Jx\| = \|x\|$, $x \in E$, $Q \in L(E, F)$ a metric surjection, if it maps the open unit ball of E onto the open unit ball of F. Examples of metric injections and metric surjections are the canonical embeddings J_M: $M \to E$, M a subspace of E, and the canonical quotient map Q_N: $F \to F/N$, N a subspace of F.

We shall moreover be interested in special classes of Banach spaces. A Banach space E possesses the metric lifting property if for all $\varepsilon > 0$ and spaces F, F_0, metric surjections $Q \in L(F, F_0)$ and operators $T \in L(E, F_0)$ there exists a lifting $\tilde{T} \in L(E, F)$ with $Q\tilde{T} = T$ and $\|\tilde{T}\| \leq (1 + \varepsilon)\|T\|$. A Banach space F possesses the metric extension property if for all Banach spaces E, E_0, metric injections $J \in L(E_0, E)$ and $T \in L(E_0, F)$ there exists an extension $\tilde{T} \in L(E, F)$ with $\tilde{T}J = T$ and $\|\tilde{T}\| = \|T\|$. For results concerning these spaces we refer to Lacey (1974) and Pietsch (1978). Examples are given later in the applications.

For a fixed Banach space E, B_E denotes the closed unit ball, i.e., $B_E = \{x \in E : \|x\| \leq 1\}$, ∂B_E its boundary.

A mapping N from B_E to some Banach space G is called information. Having fixed some information N: $B_E \to G$, any mapping from $N(B_E) \subseteq G$ to some Banach space F is called an algorithm (using information N).

Now, fix a linear operator $S \in L(E, F)$, an information operator N from B_E to some G, and an algorithm φ from $N(B_E)$ to F. We are interested in the worst-case error

$$e(S, N, \varphi) := \sup\{\|Sx - \varphi(Nx)\|, x \in B_E\}.$$

Note that instead of taking arbitrary sets $J_0 \subseteq E$, as in GTOA, we restrict ourselves to the unit ball. The problem we are concerned with is to minimize $e(S, N, \varphi)$ if N and φ vary in special classes. Let $\mathcal{B}$ denote the class of all Banach spaces.

For $n \in \mathbb{N}$, E, $G \in \mathcal{B}$ define

$$\mathfrak{N}^n_{\mathrm{arb}}(E, G) := \{N: B_E \to G, \dim \operatorname{span} N(B_E) \leq n\};$$

$$\mathfrak{N}^n_{\mathrm{con}}(E, G) := \{N: B_E \to G, N \text{ continuous}\} \cap \mathfrak{N}^n_{\mathrm{arb}}(E, G);$$

$$\mathfrak{N}^n_{\mathrm{lin}}(E, G) := \{N: B_E \to G, \text{ there is a linear extension to } E\} \cap \mathfrak{N}^n_{\mathrm{con}}(E, G);$$

$$\mathfrak{N}_{\mathrm{arb}} := \bigcup \{\mathfrak{N}^n_{\mathrm{arb}}(E, G), E, G \in \mathcal{B}, n \in \mathbb{N}\};$$

$$\mathfrak{N}_{\mathrm{con}} := \bigcup \{\mathfrak{N}^n_{\mathrm{con}}(E, G), E, G \in \mathcal{B}, n \in \mathbb{N}\};$$

$$\mathfrak{N}_{\mathrm{lin}} := \bigcup \{\mathfrak{N}^n_{\mathrm{lin}}(E, G), E, G \in \mathcal{B}, n \in \mathbb{N}\}.$$

These will be called the classes of arbitrary, continuous, and continuous linear information operators. If $\mathfrak{N} \subseteq \mathfrak{N}_{\mathrm{arb}}$ is any subclass, define for $n \in \mathbb{N}$, E, $G \in \mathcal{B}$: $\mathfrak{N}^n(E, G) := \mathfrak{N} \cap \mathfrak{N}^n_{\mathrm{arb}}(E, G)$.

With some $F \in \mathcal{B}$, $N \in \mathfrak{N}_{\mathrm{arb}}$ fixed, i.e., there are E, G, $n \in \mathbb{N}$ and $N \in \mathfrak{N}^n_{\mathrm{arb}}(E, G)$, define

$$\phi_{\mathrm{arb}}(N, F) := \{\varphi: \operatorname{Im} N \to F\};$$

$$\phi_{\mathrm{con}}(N, F) := \{\varphi: \operatorname{Im} N \to F : \varphi \text{ continuous}\};$$

$$\phi_{\mathrm{lin}}(N, F) := \{\varphi: \operatorname{Im} N \to F : \varphi \text{ admits a linear extension to } G\}.$$

Moreover, as above,

$$\phi_{\text{arb}} := \bigcup \{\phi_{\text{arb}}(N, F), N \in \mathfrak{N}_{\text{arb}}, F \in \mathcal{B}\};$$
$$\phi_{\text{con}} := \bigcup \{\phi_{\text{con}}(N, F), N \in \mathfrak{N}_{\text{arb}}, F \in \mathcal{B}\};$$
$$\phi_{\text{lin}} := \bigcup \{\phi_{\text{lin}}(N, F), N \in \mathfrak{N}_{\text{arb}}, F \in \mathcal{B}\}.$$

If $\phi \subseteq \phi_{\text{arb}}$ is any subclass, define for $N \in \mathfrak{N}_{\text{arb}}$, $F \in \mathcal{B}$:

$$\phi(N, F) := \phi \cap \phi_{\text{arb}}(N, F).$$

In GTOA it was observed that there is a deep connection between n-widths and minimal errors. Because we stress this dependence in detail, let us recall the notion of s-scales for linear operators. Denote by $\mathcal{L}$ the class of all bounded linear operators, i.e., $\mathcal{L} := \bigcup\{L(E, F), E, F \in \mathcal{B}\}$.

DEFINITION. A mapping assigning to every $S \in \mathcal{L}$ a sequence $(s_n(S))_{n\in\mathbb{N}}$ is called an s-scale, if the properties given below hold, $s_n(S)$ is called the nth s-number of S.

(S1) $\|S\| = s_1(S) \geq s_2(S) \geq \ldots \geq 0$;

(S2) $s_n(S + T) \leq s_n(S) + \|T\|$ for all $E, F \in \mathcal{B}$ and $S, T \in L(E, F)$;

(S3) $s_n(RST) \leq \|R\| s_n(S) \|T\|$ for all $E_0, E, F, F_0 \in \mathcal{B}$, $T \in L(E_0, E)$, $S \in L(E, F)$, $R \in L(F, F_0)$;

(S4) $\operatorname{rank} S < n$ implies $s_n(S) = 0$;

(S5) $s_n(\mathrm{id}: l_2^n \to l_2^n) = 1$, where id denotes the formal identity in $\mathbb{R}^n$.

The importance of this notion can be expressed in the following result.

THEOREM 1. *Let $\mathfrak{N} \subseteq \mathfrak{N}_{\text{arb}}$ be any class of information satisfying*

(N1) $\mathfrak{N}_{\text{lin}} \subseteq \mathfrak{N} \subseteq \mathfrak{N}_{\text{con}}$.

(N2) *For all $E_0, E, G \in \mathcal{B}$, $n \in \mathbb{N}$, $N \in \mathfrak{N}^n(E, G)$, $T \in L(E_0, E)$, $\|T\| \leq 1$, we have $N \circ T \in \mathfrak{N}^n(E_0, G)$.*

Let $\phi \subseteq \phi_{\text{arb}}$ be any class of algorithms satisfying

(ϕ1) $\phi_{\text{lin}} \subseteq \phi$.

(ϕ2) *For all $F_0, F \in \mathcal{B}$, $N \in \mathfrak{N}_{\text{arb}}$, $\varphi \in \phi(N, F)$, $R \in L(F, F_0)$ we have $R \circ \varphi \in \phi(N, F_0)$.*

For any $n \in \mathbb{N}$, $E, F \in \mathcal{B}$, $S \in L(E, F)$ define

$$s_n(S) := \inf\{\inf\{e(S, N, \varphi), \varphi \in \phi(N, F)\}, N \in \mathfrak{N}^{n-1}(E, G), G \in \mathcal{B}\}.$$

Then the mapping assigning to every $S \in \mathcal{L}$ the sequence $(s_n(S))_{n\in\mathbb{N}}$ forms an s-scale.

Proof. The monotonicity is clear, and, since dim span $N(B_E) = 0$ means $N = 0$, we also have $\|S\| = s_1(S)$.

To verify (S2) it is enough to observe that given $S, T \in L(E, F)$ any information operator $N \in \mathfrak{N}^n(E, G)$, then an algorithm $\varphi \in \phi(N, F)$ for S may also serve for $S + T$. To see (S3) fix $\varepsilon > 0$, R, S, T as in (S3), choose $N \in \mathfrak{N}^{n-1}(E, G)$, $\varphi \in \phi(N, F)$, such that $e(S, N, \varphi) \leq s_n(S) + \varepsilon$. Properties (N2) and ($\phi$2) imply that $\tilde{N} := N \circ (T/\|T\|) \in \mathfrak{N}^{n-1}(E_0, G)$ and $\tilde{\varphi} := \|T\|R \circ \varphi \in \phi(\tilde{N}, F)$. Moreover, $e(RST, \tilde{N}, \tilde{\varphi}) \leq \|T\|\ \|R\| e(S, N, \varphi) \leq \|T\|\ \|R\|(s_n(S) + \varepsilon)$, proving (S3).

If $S \in L(E, F)$, rank $S < n$, then $\tilde{S} = S|_{B_E}$ can be interpreted as information. Property (N1) implies that $\tilde{S} \in \mathfrak{N}^{n-1}(E, F)$. If we put $\tilde{\varphi} = \mathrm{id}_F|_{S(B_E)}$, then ($\phi$1) implies that $\tilde{\varphi} \in \phi(\tilde{S}, F)$ and we obtain $e(S, \tilde{S}, \tilde{\varphi}) = 0$, proving $s_n(S) = 0$.

Let us now turn to the verification of (S5). Obviously, we have $s_n(\mathrm{id}: l_2^n \to l_2^n) \leq 1$. On the other hand, fix $N \in \mathfrak{N}^{n-1}(l_2^n, G)$. Property (N1) implies that N is continuous. Thus we apply Borsuk's Antipodal Theorem (cf. Tichomirov, 1976, 1.7.1), to find x_0 on the boundary of $B_{l_2^n}$ with $N(x_0) = N(-x_0)$. Now we can conclude for any $\varphi \in \phi(N, l_2^n)$ that

$$e(\mathrm{id}: l_2^n \to l_2^n, N, \varphi) \geq \max\{\|x_0 - \varphi(N(x_0))\|, \|x_0 + \varphi(N(-x_0))\|\} \geq \|x_0\| = 1,$$

which ends the proof of Theorem 1.

In the spirit of Theorem 1 some remarks seem to be of interest.

Remark. 1. We did not make use of the full strength of (N1). If we denote by $\mathfrak{N}^\#$ the class of all mappings such that for $n \in \mathbb{N}$, $E, G \in \mathfrak{B}$ we have $\mathfrak{N}^{\#,n}(E, G) \subseteq \mathfrak{N}^n_{\mathrm{con}}(E, G)$ whenever dim $E < \infty$, then it follows easily that

(N1$^\#$) $\quad \mathfrak{N}_{\mathrm{lin}} \subseteq \mathfrak{N} \subseteq \mathfrak{N}^\#$

would be sufficient to apply Borsuk's Antipodal Theorem, hence proving Theorem 1 with (N1) replaced by (N1$^\#$). Roughly speaking, $\mathfrak{N}^\#$ is the class of all information operators such that the restrictions to all finite-dimensional sections of the unit ball are continuous. This class has been considered by Kacewicz and Wasilkowski (1986).

Remark 2. If the information operators are not forced to be continuous, then the above procedure may fail to produce s-scales. More precisely, we shall show that given $S \in L(E, F)$ with separable image, for all $\varepsilon > 0$ there exist $N \in \mathfrak{N}^1_{\mathrm{arb}}(E, \mathbb{R})$, $\varphi \in \phi_{\mathrm{con}}(N, F)$ with $e(S, N, \varphi) \leq \varepsilon$. Namely, let $(y_n)_{n\in\mathbb{N}}$ be a dense subset of $S(B_E) \subseteq F$. Let P be a polygon connecting the $(y_n)_{n\in\mathbb{N}}$ (in any fixed order). This gives rise to a continuous mapping Ψ from a subset of $\mathbb{R}$ to P. Take any choice Γ: $S(B_E) \to P$, such that $\|Sx - \Gamma(x)\| \leq \varepsilon$. Putting $N := \Psi^{-1} \circ \Gamma \in \mathfrak{N}^1_{\mathrm{arb}}(E, \mathbb{R})$, where Ψ^{-1} means any choice and $\varphi := \Psi$, we obtain $e(S, N, \varphi) \leq \varepsilon$.

Remark 3. If, on the other hand, $\phi = \phi_{\text{lin}}$ then the procedure of Theorem 1 produces an s-scale, the so-called Kolmogorov scale d; see GTOA, Chap. 7.4, and Pietsch (1987, 2.5.2), where the following representation is given:

$$d_n(S) = \inf\{\|Q_N S\|, \dim N < n, N \subseteq F\}.$$

EXAMPLE 1. Applying Theorem 1 with $\mathfrak{N} = \mathfrak{N}_{\text{lin}}$ and $\phi = \phi_{\text{lin}}$ we obtain the scale a of the approximation numbers. See Pietsch (1978), where this scale is introduced via

$$a_n(S) := \inf\{\|S - L\|, \operatorname{rank} L < n\}.$$

With any $S \in L(E, F)$ and information $N \in \mathfrak{N}^n_{\text{arb}}(E, G)$ chosen, the quantity $r(N, S) := \inf\{e(S, N, \varphi), \varphi \in \phi_{\text{arb}}(N, F)\}$ is a lower bound on the errors of algorithms using N. Another quantity of this kind is the diameter $d(N, S)$ of information

$$d(N, S) := \sup\{\operatorname{diam}\{Sy, Nx = Ny\}, y, x \in B_E\}$$

(cf. GTOA, Chap. 1.2).

In many cases the diameter is easier to handle and we have, for $B \subseteq F$, $\operatorname{rad}(B) \leq \operatorname{diam}(B) \leq 2 \operatorname{rad}(B)$, which yields

$$r(N, S) \leq d(N, S) \leq 2r(N, S).$$

COROLLARY 2. *Let $\mathfrak{N} \subseteq \mathfrak{N}_{\text{arb}}$ be any subclass of information, satisfying* (N1) *and* (N2). *For all $n \in \mathbb{N}$, $S \in L(E, F)$ define*

$$s_n(S) := 1/2 \inf\{d(N, S), N \in \mathfrak{N}^{n-1}(E, G), G \in \mathfrak{B}\}.$$

The mapping s assigning to every $S \in \mathfrak{L}$ the sequence $(s_n(S))_{n\in\mathbb{N}}$ constitutes an s-scale.

The proof is along the lines of the foregoing proof.

We now illustrate the corollary by two examples.

EXAMPLE 2. Putting $\mathfrak{N} = \mathfrak{N}_{\text{lin}}$ we obtain the Gelfand scale c. The following representation is well known for $n \in \mathbb{N}$ and $S \in L(E, F)$:

$$c_n(S) := \inf\{\|S|_M\|, M \subseteq E \text{ closed subspace, codim } M < n\}$$

(see GTOA, Chap. 2, Lemma 3.1).

EXAMPLE 3. Putting $\mathfrak{N} = \mathfrak{N}_{\text{con}}$ we are in a position similar to that of Babenko (1976). Let us denote the s-scale obtained by γ and call it the Babenko scale.

Concluding, in Fig. 1 we summarize the s-scales obtained for $s_n(S)$, with $n > 1$ and $S \in L(E, F)$ fixed. Note that we implicitly define some new ones.

$\mathfrak{N} \diagdown \phi$	ϕ_{arb}	ϕ_{co}	ϕ_{lin}
$\mathfrak{N}_{\text{arb}}$	$= 0$, if $S(B_E)$ separable	$= 0$, if $S(B_E)$ separable	$\underset{\sim}{d}_n(S)$
$\mathfrak{N}_{\text{con}}$	$g_n(S)$	$\tilde{g}_n(S)$	$\tilde{d}_n(S)$
$\mathfrak{N}_{\text{lin}}$	$r_n(S)$	$\tilde{r}_n(S)$	$a_n(S)$

FIGURE 1

Obviously, we have the following relations for fixed $n \in \mathbb{N}$, and $S \in \mathfrak{L}$ fixed:

(i) $1/2 g_n(S) \leq \gamma_n(S) \leq g_n(S)$,

(ii) $1/2 r_n(S) \leq c_n(S) \leq r_n(S)$,

(iii) whenever $\mathfrak{N}_1 \subseteq \mathfrak{N}_2$ and $\phi_1 \subseteq \phi_2$ we have for the associated s-scales s^1 and s^2 for every $n \in \mathbb{N}$ and $S \in \mathfrak{L} : s_n^2(S) \leq s_n^1(S)$.

2. EQUALITIES

This section is devoted to the study of various situations, where different s-numbers yield equal quantities. Well-known tools from functional analysis permit reproving several results from GTOA and obtaining new ones.

THEOREM 3. (Pietsch, 1974). *There is only one s-scale on the class of all operators acting between Hilbert spaces.*

For a proof we refer to Pietsch (1987, 2.11.9).

We shall now derive the result of Kacewicz and Wasilkowski mentioned in Section 1 from Theorems 1 and 3. Let $g^{\#}$ denote the s-scale, obtained by Theorem 1 for $\mathfrak{N} = \mathfrak{N}^{\#}$ and $\phi = \phi_{\text{arb}}$ (see also Remark 1).

COROLLARY 4. (Kacewicz and Wasilkowski, 1986). *Let S be any operator acting between Hilbert spaces. Then we have for all $n \in \mathbb{N}$ the equation $g_n^{\#}(S) = r_n(S)$.*

THEOREM 5. *Let F have the metric extension property, and let $S \in L(E, F)$, where E is an arbitrary Banach space. Then for any $N \in \mathfrak{N}_{\text{lin}}^n(E, G)$ there exists a $\varphi \in \phi_{\text{lin}}(N, F)$ with $d(N, S) = 2e(S, N, \varphi)$. Consequently, $c_n(S) = r_n(S) = a_n(S)$, $n \in \mathbb{N}$.*

Proof. Fix any information $N \in \mathfrak{N}_{\text{lin}}^n(E, G)$. the kernel ker N is a subspace in E. Since F has the metric extension property there is an $\tilde{S} \in L(E, F)$, such that, $\tilde{S}J_{\ker N} = SJ_{\ker N}$ and $\|\tilde{S}\| = \|SJ_{\ker N}\|$. Define as a linear algorithm the mapping $\varphi(Nx) := (S - \tilde{S})x$, $x \in E$. Note that ϕ is well defined and we have

$$e(S, N, \varphi) = \sup\{\|Sx - (S - \tilde{S})x\|, x \in B_E\} = \|\tilde{S}\|$$
$$= \|SJ_{\ker N}\| = 1/2d(N, S).$$

Special cases of the above theorem have been proved by Smolyak (1965), who treated the case $F = \mathbb{R}$, and Packel (1986) in terms of extending the range space.

COROLLARY 6. *Let K be a compact Hausdorff space, let E be a Banach space, and let $S \in L(E, C(K))$ be a compact operator. For all $\varepsilon >$ 0, $N \in \mathfrak{N}^n_{\text{lin}}(E, G)$ there exists a $\varphi_\varepsilon \in \phi_{\text{lin}}(N, F)$ with $2e(S, N, \varphi_\varepsilon) \leq$ $(1 + \varepsilon)d(N, S)$. Consequently, $c_n(S) = r_n(S) = a_n(S)$, $n \in \mathbb{N}$.*

Proof. Without loss of generality we may (and do) assume $1/2d(N, S)$ $= \|SJ_{\ker N}\| = 1$ and $0 \leq \varepsilon \leq 1$. Denote by J: $C(K) \rightarrow C(K)^{**}$ the canonical embedding of $C(K)$ into its bidual space $C(K)^{**}$. It is well known (cf. Lacey, 1974, Chap. 7, Sect. 21, Theorem 6) that $C(K)^{**}$ has the metric extension property. Hence, by Theorem 5 we can find an $L \in \phi_{\text{lin}}(N,$ $C(K)^{**})$ with $e(JS, N, L) = 1$.

Let $\{y_1, y_1, \ldots y_m\} \subseteq C(K)$ be the points of an $\varepsilon/4$-net of $S(B_E)$ and consider the finite-dimensional subspace $F := \text{span}\{Jy_1, Jy_2, \ldots, Jy_m\} +$ $\text{Im}\, L$ in $C(K)^{**}$. Applying the principle of local reflexivity (Lacey, 1974, Chap. 7, Sect. 23, Theorem 1) we can find an operator u: $F \rightarrow C(K)$, such that $uJy = y$, whenever $Jy \in F$ and $(1 - \varepsilon/4)\|y\| \leq \|u(y)\| \leq (1 + \varepsilon/4)\|y\|$, $y \in F$. The mapping $u \circ L$ is linear continuous into $C(K)$ and satisfies, for properly chosen y_k, the inequalities

$$\begin{aligned}
\|Sx - u \circ L \circ Nx\| &\leq \|Sx - y_k\| + \|y_k - u \circ L \circ Nx\| \\
&\leq \varepsilon/4 + \|u(Jy_k - L \circ Nx)\| \\
&\leq \varepsilon/4 + (1 + \varepsilon/4)\|Jy_k - L \circ Nx\| \\
&\leq \varepsilon/4 + (1 + \varepsilon/4)\{\|Jy_k - JSx\| + \|JSx - L \circ Nx\|\} \\
&\leq \varepsilon/4 + (1 + \varepsilon/4)(1 + \varepsilon/4) \leq 1 + \varepsilon.
\end{aligned}$$

Putting $\varphi_\varepsilon := u \circ L$, we have obtained the desired estimate.

Let us make some comments concerning the foregoing results. Corollary 6 answers a problem posed by Packel and Woźniakowski (1987, open problem 2). While Corollary 6 uses extension of compact operators, Theorem 5 makes use of such extensions for arbitrary linear operators. The latter is possible exactly, as described, for the class of all Banach spaces possessing the metric extension property, while the first extension problem leads to L_1-preduals. Both classes are studied extensively by Lindenstrauss and Tzafriri (1973, II.4.a,II.4.d). Let us mention only that a $C(K)$-

space, certainly an L_1-predual, has the metric extension property iff K is extremely disconnected, i.e., the closure of every open set in K is open again.

The situation above can be represented in the diagram

$$\begin{tikzcd} E \arrow[r, "\tilde{T}"] & F \\ E_0 \arrow[u, "J"] \arrow[ur, "T"'] & \end{tikzcd} \qquad \tilde{T}J = T.$$

The dual situation reflects the problem of lifting operators, which means

$$\begin{tikzcd} E \arrow[r, "\tilde{T}"] \arrow[dr, "T"'] & F \arrow[d, "Q"] \\ & F_0 \end{tikzcd} \qquad Q\tilde{T} = T$$

and leads to the metric lifting property, described in Section 1. It is well known, that a Banach space has the metric lifting property iff it is isometric to a suitable space $l_1(\Gamma)$, for some index set Γ (cf. Lacey, 1974, Chap. 6, Sect. 18, Theorem 9). For those spaces we have the following result.

THEOREM 7. *Let E have the metric lifting property. Let F be any Banach space and $S \in L(E, F)$. For all $\varepsilon > 0$, $N \in \mathfrak{N}^n_{\mathrm{arb}}(E, G)$, $\varphi \in \phi_{\mathrm{lin}}(N, F)$ there are $N_\varepsilon \in \mathfrak{N}^n_{\mathrm{lin}}(E, F)$ and $\varphi_\varepsilon \in \phi_{\mathrm{lin}}(F, F)$ with* $\mathrm{Im}\, \varphi_\varepsilon \circ N_\varepsilon \subseteq \mathrm{span}\, \mathrm{Im}\, \varphi \circ N$ *and $e(S, N_\varepsilon, \varphi_\varepsilon) \leq (1 + \varepsilon)e(S, N, \varphi)$, hence $d_n(S) = a_n(S)$, $n \in \mathbb{N}$.*

Proof. Let N,φ be given and $M := \mathrm{span}\, \mathrm{Im}\, \varphi \circ N(B_E)$. Let $Q\colon F \to F/M$ be the canonical quotient map. We obtain $e(S, N, \varphi) = \sup\{\|Sx - \varphi(Nx)\|, x \in B_E\} \geq \sup\{\inf\{\|Sx - y\|, y \in M\}, x \in B_E\} = \sup\{\|QSx\|, x \in B_E\} = \|QS\|$. Now we fix $\varepsilon > 0$. Using the lifting property of E, we can find $\tilde{S}$ with $Q\tilde{S} = QS$ and $\|\tilde{S}\| \leq (1 + \varepsilon)\|QS\|$. Since $\mathrm{Im}\, (\tilde{S} - S) \subseteq M$, we have $\mathrm{rank}\, (\tilde{S} - S) \leq n$. Putting as information $N_\varepsilon := (\tilde{S} - S)|_{B_E}$ and algorithm $\varphi_\varepsilon := J_M$ we see that

$$e(S, N_\varepsilon, \varphi_\varepsilon) = \|\tilde{S}\| \leq (1 + \varepsilon)\|QS\| \leq (1 + \varepsilon)e(S, N, \varphi).$$

Remark 4. Since $L_1(\mu)$ spaces for a measure μ are more common than $l_1(\Gamma)$ spaces but unfortunately do not have the metric lifting property, let us make some comments. Lacey (1974, Corollary to Theorem 8) proves that for $L_1(\mu)$ spaces there is always a lifting into the bidual, i.e., if $S \in L(L_1(\mu), F_0)$ and $Q\colon F \to F_0$ is a metric surjection then there is a lifting $\tilde{S} \in L(L_1(\mu), F_0^{**})$ with $\|\tilde{S}\| = \|S\|$. Note that for the compact operator S this lifting can be done to F and we have a situation analogous to Theorem 7 and consequently $a_n(S) = d_n(S)$, $n \in \mathbb{N}$. For details we refer to Lindenstrauss and Tzafriri (1973, Theorem II.5.26, and preceding arguments).

THEOREM 8. (GTOA, Chap. 3.4). *Let H be a Hilbert space, F be a Banach space, and $S \in L(H, F)$. For any information $N \in \mathfrak{N}^n_{\text{lin}}(H, G)$ there exists a linear central interpolatory algorithm $\varphi \in \phi_{\text{lin}}(N, F)$; consequently, $d(N, S) = 2r(N, S) = 2e(S, N, \varphi)$. This implies $r_n(S) = c_n(S) = a_n(S)$, $n \in \mathbb{N}$.*

Remark 5. Following GTOA, Chap. 1.1, we say that an algorithm $\varphi \in \phi_{\text{arb}}(N, F)$ is interpolatory for S if $\varphi(Nx) = S\tilde{x}$, for some $\tilde{x}$ with $N\tilde{x} = Nx$, and that it is central, if $\varphi(Nx) = c(x)$, where $c(x)$ is the Chebyshev center of $U(x) := \{S\tilde{x}, \tilde{x} \in B_H, N\tilde{x} = Nx\}$, provided it exists. Those algorithms enjoy good properties (see GTOA). Let us mention that $c(x)$ is certainly a center of $U(x)$ if $U(x)$ is symmetric w.r.t. $c(x)$ and central algorithms have

$$r(N, S) = e(S, N, \varphi).$$

Proof of Theorem 8. Let $N \in \mathfrak{N}^n_{\text{lin}}(H, G)$ be given and let P be the orthogonal projection onto $\ker N \subseteq H$. Consider the well-defined algorithm φ for N, given by $\varphi(Nx) := S(I - P)x$. φ is interpolatory since $N(I - P)x = Nx$. To see that φ is central, we will show that $U(x)$ is symmetric w.r.t. $S(I - P)x$. It is easy to see that $U(x) = \{Sx + Sh, x + h \in B_H, h \in \ker N\}$. So, let $Sy = Sx + Sh_0$, with $x + h_0 \in B_H$, and let $h_0 \in \ker N$ be fixed. The point symmetric to $S(I - P)x$ is

$$\begin{aligned} 2S(I - P)x - Sy &= S(2(I - P)x - (x + h_0)) \\ &= S((I - P)x - (Px + h_0)). \end{aligned}$$

It is enough to show that $(I - P)x - (Px + h_0) \in B_H$. Since $h_0 = Ph_0$, $\|x + h_0\| \leq 1$ and $(I - P)x$ is orthogonal to $P(x + h_0)$ we have $\|(I - P)x - (Px + h_0)\|^2 = \|(I - P)x - P(x + h_0)\|^2 = \|(I - P)x + P(x + h_0)\|^2 = \|x + h_0\|^2 \leq 1$; hence $r(N, S) = 1/2d(N, S) = e(S, N, \varphi)$. The rest is straightforward.

THEOREM 9. *Let K be a Hilbert space, E be a Banach space and $S \in L(E, K)$ be given. For any information $N \in \mathfrak{N}^n_{\text{arb}}(E, G)$ and any algorithm $\varphi \in \phi_{\text{lin}}(N, K)$ there are a linear continuous information $N_0 \in \mathfrak{N}^n_{\text{lin}}(E, K)$ and $\varphi_0 \in \phi_{\text{lin}}(N_0, K)$ with* $\operatorname{Im} \varphi_0 \circ N_0 \subseteq \operatorname{span} \operatorname{Im} \varphi \circ N$ *and $e(S, N_0, \varphi_0) \leq e(S, N, \varphi)$. Consequently, $d_n(S) = a_n(S)$ for all $n \in \mathbb{N}$.*

Proof. Let $M := \operatorname{span} \operatorname{Im} \varphi \circ N \subseteq K$ and let P be the orthogonal projection onto M. Then we can conclude that

$$e(S, N, \varphi) \geq \sup\{\inf\{\|Sx - y\|, y \in M\}, x \in B_E\} = \|(I - P)S\|.$$

But this lower bound is attained by putting $N_0 := PS|_{B_E}$, $\varphi_0 := J_M$.

THEOREM 10. *Let E, F be Banach spaces, $S \in L(E, F)$ be a compact operator. Then for every $\varepsilon > 0$, information $N \in \mathfrak{N}^n_{\text{arb}}(E, G)$ and algo-*

rithm $\varphi \in \phi_{\text{lin}}(N, F)$ *there are continuous information* $N_\varepsilon \in \mathfrak{N}^n_{\text{con}}(E, F)$ *and algorithm* $\varphi_\varepsilon \in \phi_{\text{lin}}(N_\varepsilon, F)$ *with* span Im $\varphi_\varepsilon \circ N_\varepsilon \subseteq$ span Im $\varphi \circ N$ *and* $e(S, N_\varepsilon, \varphi_\varepsilon) \leq (1 + \varepsilon)e(S, N, \varphi)$. *Consequently,* $\tilde{d}_n(S) = d_n(S)$.

The compactness argument used to prove Theorem 10 appears in Tichomirov (1976, 4.1.1).

Proof of Theorem 10. Fix information N and algorithm φ. Put $K :=$ span Im $\varphi \circ N$. Without loss of generality we may assume $\|Q_K S\| = 1$. Fix $1 > \varepsilon > 0$, $\delta := \varepsilon/4$. Since $S(B_E)$ was assumed to be relatively compact, we can find a finite-dimensional subspace $M \subseteq F$ and a continuous mapping Ψ: $S(B_E) \to M$ such that $\sup\{\|Sx - \Psi Sx\|, x \in B_E\} \leq \delta$ (cf. Tichomirov, 1976, 4.1.1). The subspace $F_0 := M + K$ is finite dimensional, hence admits a strictly convex norm $\|\circ\|_0$, with $\|z\| \leq \|z\|_0 \leq (1 + \delta)\|z\|$, $z \in F_0$ (cf. Tichomirov, 1976, 4.1.1). $[F_0, \|\circ\|_0]$ is strictly normed such that the metric projection P: $F_0 \to K$ is a continuous mapping with $\text{dist}_0(z, K) = \|z - Pz\|_0$, $z \in F_0$. Now put $N_\varepsilon := P\Psi S|_{B_E}$ and $\varphi_\varepsilon := J_K$ to obtain

$$\begin{aligned}
e(S, N_\varepsilon, \varphi_\varepsilon) &= \sup\{\|Sx - P\Psi Sx\|, x \in B_E\} \\
&\leq \sup\{\|Sx - \Psi Sx\| + \|\Psi Sx - P\Psi Sx\|\}, x \in B_E\} \\
&\leq \delta + \sup\{\|\Psi Sx - P\Psi Sx\|, x \in B_E\} \\
&\leq \delta + \sup\{\|\Psi Sx - P\Psi Sx\|_0, x \in B_E\} \\
&\leq \delta + \sup\{\inf\{\|\Psi Sx - y\|_0, y \in K\}, x \in B_E\} \\
&\leq \delta + (1 + \delta)\sup\{\inf\{\|\Psi Sx - y\|, y \in K\}, x \in B_E\} \\
&\leq \delta + (1 + \delta)\sup\{\inf\{\|\Psi Sx - Sx\| + \|Sx - y\|, \\
&\qquad y \in K\}, x \in B_E\} \leq \delta + (1 + \delta)(1 + \delta) \leq 1 + \varepsilon.
\end{aligned}$$

Remark 6. Results stating equality of different s-numbers in various situations are well known in approximation theory (see, e.g., Pietsch, 1987, 11.5.2, 11.5.3, 11.6.2, 11.6.3; Pinkus, 1985, Chap. II.8).

3. Inequalities

Between the various s-numbers occurring in information-based complexity there is a natural order, arising directly from the definition (see Fig. 1). On the other hand it is interesting to know what the maximal gaps between different quantities are. From the point of view, of s-number theory, this problem has been treated occasionally. We shall derive relations between s-numbers here, including examples. Usually inequalities between geometric means have been derived and it is sufficient for the consideration of sequence spaces (cf. Pietsch, 1978, 11). We claim that it would also be interesting to have precise estimates between single s-numbers.

Let us summarize the results of this section in the following theorem.

THEOREM 11. *For all* $E, F \in \mathfrak{B}$, $S \in L(E, F)$, $n \in \mathbb{N}$ *we have*

(i) $a_n(S) \leq (1 + (n - 1)^{1/2})d_n(S)$;

(ii) $a_n(S) \leq (1 + (n - 1)^{1/2})r_n(S)$;

(iii) $d_n(S) \leq n^2 g_n(S)$;

(iv) $r_n(S) \leq 2n^2 g_n(S)$;

(v) $\tilde{r}_n(S) \leq 2r_n(S)$.

Remark 7. The inequalities (i) and (ii) turn out to be optimal. Whether (iii) and (iv) are optimal is unknown. We shall, however provide an example in which the exponent is attained. Pietsch (1987, Remark 2.10.7) conjectured that this is the maximal gap. It would be of interest to find further inequalities between the quantities in Fig. 1.

Remark 8. The inequalities (i) and (ii) are well known. The proof given here follows Pietsch (1978, 11.2.2). The inequalities (iii) and (iv) are modifications of a result due to Bauhardt (1977). For a related statement see Pinkus (1985, Chap. II, Sect. 5).

Proof of Theorem 11. (i) Let $M \subseteq F$, $\dim M < n$ be fixed and such that $\|Q_M S\| \leq (1 + \varepsilon)d_n(S)$. There is a projection P onto M with $\|P\| \leq (n - 1)^{1/2}$ (cf. Pietsch, 1987, 1.5.5).

Define $N := PS|_{B_E}$, and $\varphi := J_M$, to see

$$
\begin{aligned}
a_n(S) \leq \|S - J_M PS\| &= \sup\{\|Sx - PSx\|, x \in B_E\} \\
&\leq \sup\{\inf\{\|Sx - y\| + \|PSx - y\|, y \in M\}, \\
&\quad x \in B_E\} \leq (1 + (n - 1)^{1/2})\sup\{\|Q_M Sx\|, x \in B_E\} \\
&\leq (1 + (n - 1)^{1/2})(1 + \varepsilon)d_n(S).
\end{aligned}
$$

The desired inequality follows.

(ii) Let $N \in \mathfrak{N}_{\text{lin}}^{n-1}(E, G)$ be information of rank less than n such that $\|SJ_{\ker N}\| \leq (1 + \varepsilon)c_n(S)$ (cf. Section 1, Example 2). Let P be a projection along $\ker N$ with $\|P\| \leq (n - 1)^{1/2}$ (cf. Pietsch, 1987, 1.7.17). Define $\varphi(Nx) := SPx$, to obtain

$$
\begin{aligned}
\|S - \varphi \circ N\| = \|(S(I - P)\| &\leq SJ_{\ker N}\| \, \|I - P\| \\
&\leq (1 + (n - 1)^{1/2})(1 + \varepsilon)c_n(S) \\
&\leq (1 + (n - 1)^{1/2})(1 + \varepsilon)r_n(S).
\end{aligned}
$$

(iii) The proof of (iii) is based on a result due to Bauhardt (1977). Let us, intermediately, introduce another quantity,

$$h_n(S) := \sup\{a_n(BSX), \|X: l_2 \to E\| = \|B: F \to l_2\| = 1\}.$$

This quantity is usually called the nth Hilbert number of S, and the associated s-scale the Hilbert scale. It is easy to see and it is well known that $h_n(S) \le s_n(S)$ for every s-scale s, hence $h_n(S) \le g_n(S)$. Moreover, $h_n(S) = h_n(S')$ for every $S \in L(E, F)$ (see Bauhardt, 1977). Bauhardt proved in Satz 4 that for $n \in \mathbb{N}$ and $S \in L(E, F)$ we have $d_n(S) \le n^2 h_n(S)$. The inequality (iii) is now obvious.

(iv) To prove (iv) we need additionally that $c_n(S) = d_n(S')$ (see Pietsch, 1978, 11.7.6), to conclude that

$$1/2 r_n(S) \le c_n(S) = d_n(S') \le n^2 h_n(S') = n^2 h_n(S).$$

(v) The proof of (v) involves a technique from nonlinear analysis, a continuous selection argument; see Proposition 7.2 of Michael (1956), which states the following:

If $N \in L(E, G)$ is a surjective mapping, then for every $\lambda > 1$ there exists a continuous function f: $G \to E$ with

(i) $N \circ f = \mathrm{id}_G$,
(ii) $\|f(y)\| \le \lambda \inf\{\|x\|, Nx = y\}$,
(iii) $f(\alpha y) = \alpha f(y)$, $\alpha \in \mathbb{R}$.

Let $N \in \mathfrak{N}_{\mathrm{lin}}^{n-1}(E, G)$ be a surjective information operator of rank $N < n$, and $\varepsilon > 0$. Put $\lambda = 1 + \varepsilon$. For $y \in N(B_E)$ we have $\inf\{\|x\|, Nx = y\} \le 1$, which implies $\|f(y)\| \le 1 + \varepsilon$ for all $y \in N(B_E)$. The mapping $\varphi := S \circ f$ is obviously a continuous algorithm for N, $\varphi \in \phi_{\mathrm{con}}(N, F)$.

For $x \in B_E$ we can now estimate $\|Sx - \varphi(Nx)\| = \|Sx - Sf(Nx)\| = \|S\{(x - f(Nx))/\|x - f(Nx)\|\}\| \, \|x - f(Nx)\| \le (2 + \varepsilon)\|Sh\|$, where $h = (x - f(Nx))/\|x - f(Nx)\|$ is in $(\ker N) \cap B_E$. Consequently we obtain $e(S, N, \varphi) \le (2 + \varepsilon)/2d(N, S)$, which implies $\tilde{r}_n(S) \le 2r_n(S)$.

The proof of Theorem 11 is complete.

To exhibit examples, we introduce the Sobolev spaces

$$W_p^r[0, 1] = W_p^r = \{f \in C^{r-1}[0, 1], f^{(r-1)} \text{ abs. const.}, f^{(r)} \in L_p[0, 1]\},$$

with the usual norm (see Pinkus, 1985, Chap. VII), for $1 \le p \le \infty$, $2 \le r \le \infty$, $r \in \mathbb{N}$.

EXAMPLE 4. Consider the approximation problem of W_2^r in L_∞ for $r \in \mathbb{N}$. There exist constants $0 < c \leq 1 \leq C < \infty$, such that $d_n(S) \leq Cn^{-r}$, $n \in \mathbb{N}$ (see Pinkus, 1985, Chap. VII, Theorem 3.5), while $a_n(S) \geq cn^{-r+1/2}$, $n \in \mathbb{N}$ (see Pinkus (1985, Chap. VII, Theorem 2.2), hence showing that (i) is optimal.

EXAMPLE 5. Fix $r \in \mathbb{N}$ and consider the approximation problem of W_1^r in L_2, i.e., $S = \mathrm{Id}: W_1^r \to L_2$. There exist constants $0 < c \leq 1 \leq C < \infty$, such that $c_n(S) \leq Cn^{-r}$, $n \in \mathbb{N}$ (see Pinkus (1985, Chap. VII, Theorem 3.5), while $a_n(S) \geq cn^{-r+1/2}$, $n \in \mathbb{N}$ (see Pinkus, 1985, Chap. VII, Theorem 2.2), hence (ii) is optimal.

The upper bounds in the above examples follow from a result of Kashin (1977) (see Pinkus, 1985, Chap. VI, Theorem 2.26).

EXAMPLE 6. As mentioned in Remark 7, we shall provide an example for (iii). Fix $1 < p < 2$, $m, n \in \mathbb{N}$, $m > n$. Consider the approximation problem of l_p^m in l_∞^m. Kashin (1980) has proved that $c_n(S) \geq \frac{1}{4}$, if $n \leq 0.5\, m^{2/p'}$, where $1/p + 1/p' = 1$. Now we shall follow ideas due to Babenko (1976a). Fix $E, F \in \mathfrak{B}$, $n \in \mathbb{N}$, and $S \in L(E, F)$ and consider the quantity

$$\alpha_n(S) := \inf\{\sup\{\|Sx - \phi Sx\|, x \in B_E\}, \phi \text{ odd and continuous and } \phi(S(B_E)) \text{ compact in } F \text{ with topological dimension } <n\}.$$

R. Linde (1986) has proved, that $(\alpha_n(S))_{n\in\mathbb{N}}$ forms an s-scale, the so-called Alexandrov scale (see also Tichomirov, 1976, 1.1.5; Babenko, 1976a). The Universal Space Theorem (Engelking, 1978, Theorem 1.11.4) says that there is a homeomorphism ψ: $\phi S(B_E) \to \mathbb{R}^{2n-1}$. Taking $N := \psi\phi S$ as information, we have $N \in \mathfrak{N}_{\mathrm{con}}^{2n-1}(E, \mathbb{R}^{2n-1})$. Consequently, $g_{2n}(S) \leq \alpha_n(S)$, $n \in \mathbb{N}$. But the Alexandrov numbers are well calculated; see R. Linde (1986) and Stesin (1975) for finite-dimensional embeddings: $\alpha_n(S) = (n + 1)^{-1/p}$, $n \in \mathbb{N}$ for $1 \leq p \leq 2$. Since p can be chosen arbitrarily close to 1, the exponent 1 is obtained.

EXAMPLE 7. Considering the approximation problem of l_1^m in l_q^m, $2 \leq q < \infty$, we have from Kashin (1980) that

$$d_n(S) \geq \tfrac{1}{4}, \qquad n < 0.5\, m^{2/q},$$

while

$$g_{2n}(S) \leq \alpha_n(S) \leq (n + 1)^{1/q-1}$$

(see R. Linde, 1986). Thus, (iv) of Theorem 10 cannot be proved for some exponent $0 < p < 1$.

PART B: THE AVERAGE-CASE SETTING

4. Notation and General Results

As in Part A we determine the error produced by some information and appropriate algorithm for linear problems. In the average-case setting, this error depends on the choice of some measure on the underlying Banach space. Roughly speaking, instead of taking the supremum over the unit ball we average over the whole space.

To be precise, let E be a Banach space, $\mathfrak{B}(E)$ the σ-algebra of Borel sets on E, and μ a positive Radon measure on $[E, \mathfrak{B}(E)]$, i.e., $\mu(A) = \sup\{\mu(K), K \subseteq A \text{ compact}\}$, $A \in \mathfrak{B}(E)$. Denote by $\mathfrak{M}(E)$ the set

$$\mathfrak{M}(E) := \{\mu,\ \mu \text{ is a symmetric positive Radon measure on } [E, \mathfrak{B}(E)] \text{ with } \textstyle\int \|x\| d\mu(x) < \infty\}.$$

For facts concerning Radon measures we refer to Parthasarathy (1967) and W. Linde (1983).

For natural reasons we restrict all considerations to information operators and algorithms which are Borel measurable.

Given $E, F, G \in \mathfrak{B}$, $n \in \mathbb{N}$ we define

$$\overline{\mathfrak{N}}^n_{\text{mes}}(E, G) := \{N: E \to G,\ N \text{ Borel measurable, dim span } N(E) \le n\};$$

$$\overline{\mathfrak{N}}^n_{\text{con}}(E, G) := \{N: E \to G,\ N \text{ continuous, dim span } N(E) \le n\};$$

$$\overline{\mathfrak{N}}^n_{\text{lin}}(E, G) := \{N: E \to G,\ N \text{ continuous linear, } \underline{\dim}\ N(E) \le n\}$$

and as in Section 1 we construct $\overline{\mathfrak{N}}_{\text{mes}}$, $\overline{\mathfrak{N}}_{\text{con}}$, $\overline{\mathfrak{N}}_{\text{lin}}$.

Now, having fixed $N \in \overline{\mathfrak{N}}^n_{\text{mes}}(E, G)$, we define

$$\bar{\phi}_{\text{mes}}(N, F) := \{\varphi: N(E) \to F,\ \varphi \text{ is the restriction of a measurable mapping}\};$$

$$\bar{\phi}_{\text{con}}(N, F) := \{\varphi: N(E) \to F,\ \varphi \text{ continuous}\};$$

$$\bar{\phi}_{\text{lin}}(N, F) := \{\varphi: N(E) \to F,\ \varphi \text{ linear}\}.$$

(Note that $\varphi \in \bar{\phi}_{\text{lin}}(N, F)$ is always continuous, since dim span $N(E) < \infty$.)

We proceed further, as in Section 1, to obtain $\bar{\phi}_{\text{mes}}$, $\bar{\phi}_{\text{con}}$, and $\bar{\phi}_{\text{lin}}$. With $S \in L(E, F)$ and $\mu \in \mathfrak{M}(E)$, information $N \in \overline{\mathfrak{N}}^n_{\text{mes}}(E, G)$, and an algorithm $\varphi \in \bar{\phi}_{\text{mes}}(N, F)$ chosen, the average-case error $e(S, N, \varphi, \mu)$ is defined to be $e(S, N, \varphi, \mu) = \int \|Sx - \varphi(N(x))\| d\mu(x)$, called the μ-error for the problem S using N and φ.

As in Part A we are interested in minimizing the μ-error, if N and φ vary in special classes.

In a comparison of the situations in the worst-case and average-case settings one of the main results of Kacewicz and Wasilkowski (1986, Theorem 3.1) says that for Hilbert spaces H, K, and $S \in L(H, K)$ and $\mu \in \mathfrak{M}(H)$ the quantity $\inf\{e(S, N, \varphi, \mu), N \in \overline{\mathfrak{N}}^n_{\text{con}}(H, \mathbb{R}), \varphi \in \overline{\phi}_{\text{mes}}(N, K)\}$ is equal to zero.

The first theorem of this section is fundamental to generalizing this fact (see Corollary 13).

THEOREM 12. Let E be a Banach space.

(i) For any positive Radon measure μ on E, $\mu(E) = 1$ and every $\varepsilon > 0$ there exist a continuous mapping $N \in \overline{\mathfrak{N}}^1_{\text{con}}(E, \mathbb{R})$ and a continuous mapping $\varphi \in \overline{\phi}_{\text{con}}(N, E)$, such that $\mu(\{x \in E: \|x - \varphi(N(x))\| > \varepsilon\}) \leq \varepsilon$.

(ii) For every $p > 0$, every positive Radon measure μ on E, with $\int\|x\|^p d\mu(x) < \infty$, every $\varepsilon > 0$, there exist a continuous mapping $N \in \overline{\mathfrak{N}}^1_{\text{con}}(E, \mathbb{R})$ and a continuous mapping $\varphi \in \overline{\phi}_{\text{con}}(N, E)$, such that

$$\int \|x - \varphi(N(x))\|^p d\mu(x) \leq \varepsilon.$$

Proof. We give only the proof of (i) and indicate the modifications needed for (ii), unless they are obvious.

Fix $\varepsilon > 0$. For $R > 0$ let $B_R := \{x \in E: \|x\| < R\}$. Given a set A, let A^C be its complement. We can find $R > 0$, such that $\mu(B_R^C) \leq \varepsilon/4$ (for (ii), $\int_{B_R^C}\|x\|^p d\mu(x) \leq \varepsilon/4$).

Since μ is Radon, there exists a compact set $K \subseteq B_R$ with $\mu(B_R \backslash K) \leq \varepsilon/4$ (for (ii), $\mu(B_R\backslash K) \leq \varepsilon/4(2R)^{-p}$). Now, there exists a finite $\varepsilon/8$-covering $\tilde{U}_1, \ldots, \tilde{U}_m$ of K, consisting of open sets with radius less than $\varepsilon/8$. Let $U_j := \tilde{U}_j \cap B_R$, $j := 1, \ldots, m$. Successively, since μ is Radon, we can find disjoint closed sets F_j in U_j with $\mu(U_j \backslash \bigcup_{i=1}^{j} F_i) \leq \varepsilon/4m$ (for (ii), $\mu(U_j\backslash\bigcup_{i=1}^{j}F_i) \leq \varepsilon/4m(2R)^{-p}$), $j = 1, \ldots, m$. Consider the closed set $F := \bigcup_{i=1}^{m} F_i \cup B_R^C$.

Let us define a continuous mapping $N_0: F \to \mathbb{R}$ by

$$N_0|_{F_j} := j, \qquad j = 1, \ldots, m;$$
$$N_0|_{B_R^C} := 0.$$

By the Tietze–Dugundji Extension Theorem (cf. Bessaga and Pełczynski, 1975, Chap. II, Theorem 3.1), there exists a continuous extension N onto E with $N(E) \subseteq [0, m]$. Now, choose $x_j \in B_R$, such that $\sup\{\|x - x_j\|, x \in F_j\} \leq \varepsilon/4, j = 1, \ldots, m$, and define a continuous mapping $\varphi_0: \{0, 1, \ldots, m\} \to E$ via

$$\varphi_0(j) := x_j, \qquad j = 1, \ldots, m;$$
$$\varphi_0(0) := 0.$$

Applying the Tietze–Dugundji Extension Theorem once more, we obtain a continuous $\varphi\colon \mathbb{R} \to B_R$, hence $N \in \overline{\mathfrak{N}}^1_{\text{con}}(E, \mathbb{R})$ and $\varphi \in \overline{\Phi}_{\text{con}}(N, E)$. Since $K \subseteq B_R$ implies

$$(B_R \backslash K) \cup \left(K \Big\backslash \bigcup_{i=1}^m F_i\right) = B_R \backslash \left(K \cap \bigcup_{i=1}^m F_i\right),$$

we have the estimates

$$\mu(\{x \in E : \|x - \varphi(N(x))\| > \varepsilon\}) \leq \mu\left(\left(\bigcup_{i=1}^m F_i\right)^{\mathrm{C}}\right)$$
$$\leq \mu(B_R^{\mathrm{C}}) + \mu\left(B_R \Big\backslash \bigcup_{i=1}^m F_i\right) \leq \varepsilon/4 + \mu\left(B_R \Big\backslash \left(K \cap \bigcup_{i=1}^m F_i\right)\right)$$
$$\leq \varepsilon/4 + \mu(B_R \backslash K) + \mu\left(K \Big\backslash \bigcup_{i=1}^m F_i\right)$$
$$\leq \varepsilon/2 + \sum_{j=1}^m \mu\left(U_j \Big\backslash \bigcup_{i=1}^j F_i\right) \leq \varepsilon.$$

One can use the same decomposition to get (ii).

From Theorem 12 we immediately obtain

COROLLARY 13. *For $E, F, \in \mathfrak{B}$, $S \in L(E, F)$, $\mu \in \mathfrak{M}(E)$, we have*

$$\inf\{e(S, N, \varphi, \mu), \varphi \in \overline{\Phi}_{\text{con}}(N, F), N \in \overline{\mathfrak{N}}^1_{\text{con}}(E, \mathbb{R})\} = 0.$$

For later use, let us prove the following lemma.

LEMMA 14. *Let $E, F, G \in \mathfrak{B}$, $S \in L(E, F)$, $\mu \in \mathfrak{M}(E)$, $n \in \mathbb{N}$ be given. For every $\varepsilon > 0$, $N \in \overline{\mathfrak{N}}^n_{\text{mes}}(E, G)$ and $\varphi \in \overline{\Phi}_{\text{mes}}(N, F)$ there is a continuous algorithm $\varphi_c \in \overline{\Phi}_{\text{con}}(N, F)$ with*

$$e(S, N, \varphi_c, \mu) \leq \varepsilon + e(S, N, \varphi, \mu).$$

Proof. Consider the image measure $N(\mu)$ of μ under N on G. Without loss of generality we may assume that $e(S, N, \varphi, \mu) < \infty$, hence $\int \|\varphi(Nx))\| d\mu(x) < \infty$, i.e., $\varphi \in L_1(N(\mu), F)$. Since continuous functions are dense in $L_1(N(\mu), F)$, we can find a φ_c with $\int \|\varphi(Nx) - \varphi_c(Nx)\| d\mu(x) \leq \varepsilon$, yielding

$$
\begin{aligned}
e(S, N, \varphi_c, \mu) &= \int \|Sx - \varphi_c(Nx)\| d\mu(x) \\
&\leq \int \|Sx - \varphi(Nx)\| d\mu(x) + \int \|\varphi(Nx) - \varphi_c(Nx)\| d\mu(x) \\
&\leq e(S, N, \varphi, \mu) + \varepsilon.
\end{aligned}
$$

LEMMA 15. *Let $S \in L(E, F)$ and $\mu \in \mathfrak{M}(E)$ be given. For every $\varepsilon > 0$, $N \in \overline{\mathfrak{N}}^n_{\mathrm{mes}}(E, G)$ and $\varphi \in \bar{\phi}_{\mathrm{lin}}(N, F)$ there are a continuous $N_\varepsilon \in \overline{\mathfrak{N}}_{\mathrm{con}}(E, F)$ and $\varphi_\varepsilon \in \bar{\phi}_{\mathrm{lin}}(N_\varepsilon, F)$ with* $\operatorname{Im} \varphi_\varepsilon \circ N_\varepsilon \subseteq \operatorname{span} \operatorname{Im} \varphi \circ N$ *and $e(S, N_\varepsilon, \varphi_\varepsilon, \mu) \leq e(S, N, \varphi, \mu) + \varepsilon$.*

Proof. The arguments follow the lines of the proof of Theorem 10 of Part A, after the considerations have been restricted to a compact set with almost full measure.

In view of Part A we define the following quantities.

DEFINITION. Let $S \in L(E, F)$, $\mu \in \mathfrak{M}(E)$, $n \in \mathbb{N}$ be fixed.

(i) $c_n(S, \mu) := \inf\{e(S, N, \varphi, \mu), N \in \overline{\mathfrak{N}}^{n-1}_{\mathrm{lin}}(E, G), \varphi \in \bar{\phi}_{\mathrm{mes}}(N, F), G \in \mathfrak{B}\}$ is said to be the nth μ-Gelfand number of S.

(ii) $a_n(S, \mu) := \inf\{e(S, N, \varphi, \mu), N \in \overline{\mathfrak{N}}^{n-1}_{\mathrm{lin}}(E, G), \varphi \in \bar{\phi}_{\mathrm{lin}}(N, F), G \in \mathfrak{B}\}$ is said to be the nth μ-approximation number of S.

(iii) $d_n(S, \mu) := \inf\{e(S, N, \varphi, \mu), N \in \overline{\mathfrak{N}}^{n-1}_{\mathrm{mes}}(E, G), \varphi \in \bar{\phi}_{\mathrm{lin}}(N, F), G \in \mathfrak{B}\}$ is said to be the nth μ-Kolmogorov number of S.

(C) The mapping c^{ave}, assigning to every $S \in L(E, F)$, $\mu \in \mathfrak{M}(E)$ the sequence $(c_n(S, \mu))_{n \in \mathbb{N}}$ is said to be the average Gelfand scale.

(A) The mapping a^{ave}, assigning to every $S \in L(E, F)$, $\mu \in \mathfrak{M}(E)$ the sequence $(a_n(S, \mu))_{n \in \mathbb{N}}$ is said to be the average approximation scale.

(D) The mapping d^{ave}, assigning to every $S \in L(E, F)$, $\mu \in \mathfrak{M}(E)$ the sequence $(d_n(S, \mu))_{n \in \mathbb{N}}$ is said to be the average Kolmogorov scale.

In the spirit of Part A and using the results proved before, we obtain the scheme shown in Fig. 2 for fixed $S \in L(E, F)$, $\mu \in \mathfrak{M}(E)$, and $n > 1$.

Remark 9. The average Gelfand scale appears implicitly in most papers on average-case setting, while the average Kolmogorov scale has been proposed by Micchelli (1984).

$\bar{\mathfrak{N}}$ \ $\bar{\phi}$	$\bar{\phi}_{\mathrm{mes}}$	$\bar{\phi}_{\mathrm{con}}$	$\bar{\phi}_{\mathrm{lin}}$
$\bar{\mathfrak{N}}_{\mathrm{mes}}$	$= 0$	$= 0$	$d_n(S, \mu)$
$\bar{\mathfrak{N}}_{\mathrm{con}}$	$= 0$	$= 0$	$d_n(S, \mu)$
$\bar{\mathfrak{N}}_{\mathrm{lin}}$	$c_n(S, \mu)$	$c_n(S, \mu)$	$a_n(S, \mu)$

FIGURE 2

Let us mention the obvious representations

$$a_n(S, \mu) = \inf\{\int \|Sx - Lx\| d\mu(x), \text{ rank } L < n\},$$

$d_n(S, \mu) = \inf\{\int \|QSx\| d\mu(x)$, Q quotient map over some at most $(n - 1)$-dimensional subspace of $F\} = \inf\{\int \operatorname{dist}(Sx, M) d\mu(x), \dim M < n\}$.

Remark 10. Though not having introduced the notion of an average s-scale, let us state some obvious properties of the quantities obtained. Let $s^{\text{ave}} \in \{c^{\text{ave}}, a^{\text{ave}}, d^{\text{ave}}\}$. Then we have

(i) $s_1(S, \mu) = \int \|Sx\| d\mu(x) \geq s_2(S, \mu) \geq \ldots \geq 0$; $S \in L(E, F)$, $\mu \in \mathfrak{M}(E)$;

(ii) $s_n(S + T, \mu) \leq s_n(S, \mu) + s_1(T, \mu)$ for $S, T \in L(E, F)$, $\mu \in \mathfrak{M}(E)$;

(iii) $s_n(S, \mu) = 0$, if $S \in L(E, F)$ of rank $S < n$ and $\mu \in \mathfrak{M}(E)$;

(iv) $s_n(RST, \mu) \leq \|R\| s_n(S, T(\mu))$, for $R \in L(F, F_0)$, $S \in L(E, F)$, $T \in L(E_0, E)$, and $\mu \in \mathfrak{M}(E_0)$.

Moreover, the average approximation scale has the following property: if any average s-scale s satisfies (i), (ii) and (iii) above then $s_n(S, \mu) \leq a_n(S, \mu)$ for all $n \in \mathbb{N}$ and S, μ.

It would be of a great interest to the author to know what (natural) property should be added, to obtain unicity of average s-scale on the class of all operators $S \in L(E, K)$, where K is a Hilbert space and measures $\mu \in \mathfrak{M}(E)$ are Gaussian.

The following result relates the usual s-scales to their average counterparts. The average Gelfand scale seems to play a special role.

THEOREM 16. For $S \in L(E, F)$, $\mu \in \mathfrak{M}(E)$, $n \in \mathbb{N}$ *we have*

(i) $a_n(S, \mu) \leq a_n(S) \int \|x\| d\mu(x)$,

(ii) $d_n(S, \mu) \leq d_n(S) \int \|x\| d\mu(x)$.

Proof. The result follows immediately from the representations given in Remark 9.

We will not introduce the notion of a Gaussian measure in $\mathfrak{M}(E)$. The interested reader is referred to Kuo (1975) and W. Linde (1983). The canonical Gaussian measure on $\mathbb{R}^m$ (W. Linde, 1983, Example in 4.3) and $S = \mathrm{id}: \mathbb{R}^m \to l_2^m$ show that these bounds in Theorem 16 cannot be improved essentially.

In Part A we did not consider the problem whether s-numbers must tend to zero. For information see Pietsch (1978, 14.1.13) and Pinkus (1985, Chap. II, Propositions 7.1, 7.4, 7.5). Theorem 17 adapts those results to the average-case setting.

Recall that a Banach space E has the bounded approximation property (abbreviated as b.a.p.) if there exists a positive λ, such that for every $\varepsilon > 0$, compact $K \subseteq E$ there is a finite rank operator $L \in L(E, E)$ with $\|L\| \leq \lambda$ and $\sup\{\|x - Lx\|, x \in K\} \leq \varepsilon$. Though most of the common Banach spaces share this property, there are prominent counterexamples. For more information see Pietsch (1978, 10) and the references given therein. Below we shall see that spaces possessing the bounded approximation property allow linear approximation of finite rank of arbitrarily small error for every measure in $\mathfrak{M}(E)$.

It is not clear if the assumption of b.a.p. is needed. More explicitly, is the following true? Given any Banach space E and measure $\mu \in \mathfrak{M}(E)$, does $(a_n(S, \mu))_{n\in\mathbb{N}}$ converge to zero. Let us denote by c_0 the space of all sequences converging to zero. Remember that we have $c_n(S, \mu) \leq a_n(S, \mu)$ directly from the definition.

THEOREM 17. *Let $E, F \in \mathfrak{B}$, $S \in L(E, F)$, $\mu \in (E)$.*

(i) If E has the b.a.p., then $(a_n(S, \mu))_{n\in\mathbb{N}} \in c_0$.

(ii) $(d_n(S, \mu))_{n\in\mathbb{N}} \in c_0$.

Proof. It is enough to prove (i) and (ii) for $S := \mathrm{id}_E$. So, let $\varepsilon > 0$, let $K \subseteq E$ be compact with $\int_{K^c}\|x\|d\mu(x) \leq \varepsilon$. Using the b.a.p. we can find $\lambda > 0$ and $L \in L(E, E)$ of finite rank with $\|L\| \leq \lambda$ and $\sup\{\|x - Lx\|, x \in K\} \leq \varepsilon$. Now we can conclude that

$$\begin{aligned}\int\|x - Lx\|d\mu(x) &= \int_K\|x - Lx\|d\mu(x) + \int_{K^c}\|x - Lx\|d\mu(x)\\ &\leq \varepsilon + \varepsilon(1 + \lambda) \leq 1\varepsilon(1 + \lambda).\end{aligned}$$

If we let $\varepsilon \to 0$, (i) follows.

To prove (ii) let K be compact with $\int_{K^c}\|x\|d\mu(x) \leq \varepsilon$. Further, put $\{x_1, x_3, \ldots, x_m\}$ as the points of an ε-net of $K \subseteq F$ and $M := \mathrm{span}\{x_1, x_2, \ldots, x_m\}$. We obtain

$$\begin{aligned}d_{m+1}(\mathrm{id}_E, \mu) &\leq \int \mathrm{dist}(x, M)d\mu(x)\\ &= \int_K \mathrm{dist}\,(x, M)d\mu(x) + \int_{K^c} \mathrm{dist}(x, M)d\mu(x) \leq 2\varepsilon,\end{aligned}$$

yielding (ii).

To point out the distinct behavior in the average-case and worst-case settings we shall need the following lemma. A measure $\mu \in \mathfrak{M}(E)$ is said to be finitely supported, if it admits a representation $\mu = \sum_{i=1}^m \lambda_i(\delta_{x_i} + \delta_{-x_i})$ for some $m \in \mathbb{N}$, $(\lambda_k)_{k=1}^m$, and $\{x_1, x_2, \ldots, x_m\} \subseteq E$. (Given $x \in E$ we denote by δ_x the Dirac measure at x, i.e.,

$$\delta_x(A) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases} \qquad A \in \mathfrak{B}(E).)$$

LEMMA 18. *Let $\mu \in \mathfrak{M}(E)$ be finitely supported. Then there exists a functional $a \in E'$ and a continuous mapping ψ: $\mathbb{R} \to E$, such that* $\mathrm{id}_E = \psi \circ a\ \mu$ – a.s.

Proof. Let $\mu = \sum_{i=1}^{m} \lambda_i(\delta_{x_i} + \delta_{-x_i})$ be finitely supported with distinct $\{x_1, x_2, \ldots, x_m\} \subseteq E$. Put $y_{i,j} := x_i - x_j$, $1 \le i < j \le m$, hence $y_{i,j} \ne 0$. Let $A_{i,j} := \{a \in E', \langle y_{i,j}, a\rangle = 0\}$. $A_{i,j}$ has codimension 1 in E', hence has empty interior. Consequently, $E' \ne \bigcup_{1 \le i < j \le m} A_{i,j}$, so there is $a \in E'$ not belonging to $A_{i,j}$, $1 \le i < j \le m$. This means that $\langle y_{i,j}, a\rangle \ne 0$ for all $1 \le i < j < m$. Hence a is injective on $\{x_1, x_2, \ldots, x_m\}$, taking different values r_k at x_k, $k = 1, \ldots, m$. Any continuous extension ψ of $\psi|_{\{r_j\}} = x_j$, $j = 1, \ldots, m$, may serve to show $\psi \circ a = \mathrm{id}_E\ \mu$ – a.s.

COROLLARY 19. *Let $S \in L(E, F)$ and $\mu \in \mathfrak{M}(E)$ be finitely supported. Then there exist $N \in \overline{\mathfrak{N}}^1_{\mathrm{lin}}(E, \mathbb{R})$ and $\varphi \in \overline{\phi}_{\mathrm{con}}(N, F)$ with $e(S, N, \varphi, \mu) = 0$. Consequently, $c_2(S, \mu) = 0$.*

Proof. Choose a, ψ from Lemma 18 and put $N := a$ and $\varphi := S \circ \psi$.

EXAMPLE 8. Consider $E = \mathbb{R}^2$, $S = \mathrm{id}_E$ and $\mu = \delta_{(2,2)} + \delta_{(-2,-2)} + \delta_{(1,-1)} + \delta_{(-1,1)}$. The information $N(x, y) = x$ admits μ-a.s. a continuous inverse φ which is defined by $\varphi(r) = (r, 1/3r(2r^2 - 5))$.

Remark 11. If dim span(supp(μ)) > 1, then $d_2(\mathrm{id}_E, \mu)$ cannot be equal to 0. Moreover, let $\mu \in \mathfrak{M}(E)$ be finitely supported and let $\nu \in \mathfrak{M}(E)$ be Gaussian, with its support containing supp(μ). It is well known that the support of ν is a linear subspace in E (cf. W. Linde, 1983, 6.9). Thus, the measure $\bar{\mu} = (1 - \varepsilon)\mu + \varepsilon\nu$ belongs to $\mathfrak{M}(E)$ and has a linear support. Now choose $N \in \overline{\mathfrak{N}}^1_{\mathrm{lin}}(E, \mathbb{R})$, $\varphi \in \overline{\phi}_{\mathrm{con}}(N, F)$ for $S \in L(E, F)$ with $e(S, N, \varphi, \mu) = 0$. This implies that $c_2(S, \bar{\mu}) \le e(S, N, \varphi, \bar{\mu}) \le \varepsilon\int \|Sx \to \varphi(Nx)\|\, d\nu(x)$, hence $c_2(S, \bar{\mu})$ can be made as small as we want by the proper choice of $\varepsilon \ge 0$. On the other hand, $d_2(S, \bar{\mu}) \ge (1 - \varepsilon)d_2(S, \mu)$. Summarizing, we have constructed a measure $\bar{\mu} \in \mathfrak{M}(E)$ with "nice" support, still making the gap between $c_2(S, \bar{\mu})$ and $d_2(S, \bar{\mu})$ as large as we want.

Remark 12. In contrast to the worst case, Corollary 19 implies that (iii) of Remark 10 is not an equivalence.

5. EQUALITIES

Following the lines of the worst-case setting we shall study the assumptions under which the mappings c^{ave}, a^{ave}, and d^{ave} may be equal. The first result is an immediate consequence of the ideas presented in the proof of Theorem 9 of Part A.

THEOREM 20. *Let K be a Hilbert space, $E \in \mathfrak{B}$, $S \in L(E, K)$, $n \in \mathbb{N}$, $\mu \in \mathfrak{M}(E)$. For any information $N \in \overline{\mathfrak{N}}^{n}_{\mathrm{mes}}(E, G)$ and algorithm $\varphi \in \overline{\phi}_{\mathrm{lin}}(N, F)$ there are a continuous linear information $N_0 \in \overline{\mathfrak{N}}^{n}_{\mathrm{lin}}(E, F)$ and $\varphi_0 \in \overline{\phi}_{\mathrm{lin}}(N_0, F)$ with* $\mathrm{Im}\ \varphi_0 \circ N_0 \subseteq$ *span Im $\varphi \circ N$ and $e(S, N_0, \varphi_0, \mu) \leq e(S, N, \varphi, \mu)$. Consequently, $d_n(S, \mu) = a_n(S, \mu)$, $n \in \mathbb{N}$.*

Thus, the situations in the average-case and worst-case settings are quite similar w.r.t. the Kolmogorov and approximation scales. Remark 11 showed that the average Gelfand scale is most sensitive w.r.t. the underlying measure. Thus the equality of the average Gelfand scale and the average approximation scale depends on the measure instead of the spaces. First results stating such equalities have been proved by Lee and Wasilkowski (1986) for Gaussian measures and Wasilkowski and Woźniakowski (1984) for so-called orthogonally invariant measures.

At this point let us turn to a property which is shared by all the classes of measures mentioned above and which is easy to handle.

DEFINITION. A measure $\mu \in \mathfrak{M}(E)$ is said to be reflectable if for all finite codimensional subspaces E_0 there is an $E_1 \subseteq E$ such that, $E = E_0 \oplus E_1$ is the direct sum and μ is invariant w.r.t. $x_0 + x_1 \to -x_0 + x_1$, a reflection at E_1 along E_0.

Before giving the application, let us give an example.

EXAMPLE 9. (Orthogonally invariant measures; Wasilkowski and Woźniakowski, 1984). A measure μ on a Hilbert space H,$\int\|x\|^2 d\mu(x) < \infty$, is called orthogonally invariant if it is symmetric and invariant w.r.t. all mappings I-2P, where P is a projection of rank $P = 1$, orthogonal w.r.t. the semiscalar product on H, induced by $(f, g)_\mu := \int(f, x)(g, x)d\mu(x)$. This means that P admits a representation $P = f \otimes g$ with $(f, g)_\mu = 1$.

We claim that orthogonally invariant measures are reflectable. To see this, observe that $(I - 2P_1)(I - 2P_2) \circ\circ\circ (I - 2P_k) = I - 2\Sigma_{i=1}^{k} P_i$, whenever $\{P_1, P_2, \ldots, P_k\}$ are mutually orthogonal. thus orthogonal invariance implies the invariance under all reflections $I - 2P$, where P is of finite rank and orthogonal w.r.t. $(\ ,\ )_\mu$. The symmetry of μ implies the invariance under all reflections of finite defect. Now, μ is reflectable with $E_1 := E_0^{\perp}$ in $[E, (\ ,\ )_\mu]$. The advantage of reflectability is that we can treat also measures without second moments and need not introduce a Hilbert space. For instance, sub-Gaussian p-stable measures, i.e., p-stable measures, generated by Gaussian ones (see W. Linde, 1983, 7.6), are easily seen to be reflectable. On the other hand, W. Linde and Mathé (1983) constructed a symmetric stable measure on $\mathbb{R}^2$, which is not reflectable.

THEOREM 21. *Let $\mu \in \mathfrak{M}(E)$ be reflectable and $S \in L(E, F)$. Let $N \in \overline{\mathfrak{N}}^{n}_{\mathrm{lin}}(E, G)$ be any information and P be the projection onto* ker N, *such* *that μ is invariant* w.r.t. $I - 2P$. *For all algorithms $\varphi \in \bar{\Phi}_{\mathrm{mes}}(N, F)$ we have $e(S, N, \varphi, \mu) \geq \int \|SPx\| d\mu(x)$, hence $c_n(S, \mu) = a_n(S, \mu)$, $n \in \mathbb{N}$.*

Proof. Fix S, N, φ, μ, and P as above and observe that $NPx = 0$, $x \in E$. Moreover $P(I - 2P) = -P$ and $(I - P)(I - 2P) = I - P$ and the reflectability of μ implies

$$\begin{aligned} e(S, N, \varphi, \mu) &= \int \|Sx - \varphi(Nx)\| d\mu(x) \\ &= \int \|SPx + S(I - P)x - \varphi(N(Px + (I - P)x))\| d\mu(x) \\ &= \int \|SPx + \{S(I - P)x - \varphi(N(I - P)x))\}\| d\mu(x) \\ &= \int \|-SPx + \{S(I - P)x - \varphi(N(I - P)x))\}\| d\mu(x) \\ &= \int \|SPx - \{S(I - P)x - \varphi(N(I - P)x))\}\| d\mu(x). \end{aligned}$$

The convexity of the norm yields that $\int \|SPx\| d\mu(x)$ is a lower bound for $e(S, N, \varphi, \mu)$. This bound is attained by the (well-defined) linear algorithm $\varphi_{\mathrm{L}}(Nx) = S(I - P)x$, $x \in E$.

6. Inequalities

Here we look for the maximal gap between c^{ave}, a^{ave}, and d^{ave}. Corollary 19 and Remark 11 of Section 4 imply that there does not exist a (finite) function f: $\mathbb{N} \to \mathbb{R}^+$, such that for all $S \in L(E, F)$, $\mu \in \mathfrak{M}(E)$ we have $d_n(S, \mu) \leq f(n)c_n(S, \mu)$ or $a_n(S, \mu) \leq f(n)c_n(S, \mu)$. This contrasts to the worst-case; see Section 3, Theorem 11. On the other hand one can easily prove the next result.

THEOREM 22. *For $E, F \in \mathfrak{B}$, $S \in L(E, F)$, $\mu \in \mathfrak{M}(E)$, and $n \in \mathbb{N}$ we have*

$$a_n(S, \mu) \leq (1 + (n - 1)^{1/2}) d_n(S, \mu).$$

The proof is along the same lines as (i) of Theorem 11 in Section 3 and we omit it. It would be of great interest to establish an example where this gap is attained.

References

BABENKO, K. I. (1976a), Estimating the quality of computational Algorithms, 1, *Comput. Methods Appl. Mech. Engrg.* **7**, 47–63.

BABENKO, K. I. (1976b), Estimating the quality of computational Algorithms, 2, *Comput. Methods Appl. Mech. Engrg.* **7**, 135–152.

BAUHARDT, W. (1977), Hilbert-Zahlen von Operatoren in Banachräumen, *Math. Nachr.* **79,** 181–187.

BESSAGA, C., AND PELCZYNSKI, A. (1975), "Selected Topics in Infinite-Dimensional Topology," Polish Sci. Pub., Warsaw.

ENGELKING, R. (1978), "Dimension Theory," Polish Sci. Pub., Warsaw.

KACEWICZ, B., AND WASILKOWSKI, G. W. (1986), How powerful is continuous nonlinear information for linear problems? *J. Complexity* **2,** 306–316.

KASHIN, B. S. (1977), Diameters of some finite-dimensional sets and of some classes of smooth functions, *Izv. Akad. Nauk SSSR Ser. Mat.* **41,** 334–351.

KASHIN, B. S. (1980), On a property of Bilinear forms, *Notes Acad. Sci. Georg. SSR* **97,** 29–32.

KUO, H. H. (1975), "Gaussian Measures in Banach Spaces," Lecture Notes in Mathematics, Vol. 463, Springer-Verlag, Berlin/Heidelberg/New York.

LACEY, H. E. (1974), "The Isometric Theory of Classical Banach Spaces," Springer-Verlag, Berlin/Heidelberg/New York.

LEE, D., AND WASILKOWSKI, G. W. (1986), Approximation of linear functionals on a Banach space with a Gaussian measure, *J. Complexity* **2,** 12–43

LINDE, R. (1986), About Alexandrov-Numbers, *Forschungserg. FSU* N/86/34, Jena, unpublished.

LINDE, W. (1983), "Infinitely Divisible and Stable Measures on Banach Spaces," Teubner-Verlag, Leipzig.

LINDE, W., AND MATHÉ, P. (1983), Conditional symmetries of stable measures on $\mathbb{R}^n$, *Ann. Inst. H. Poincaré (B)*, 57–69.

LINDENSTRAUSS, J., AND TZAFRIRI, L. (1973), "Classical Banach Spaces," Lecture Notes in Mathematics, Vol. 338, Springer-Verlag, Berlin/Heidelberg/New York.

MICCHELLI, C. A. (1984), Orthogonal projections are optimal algorithms, *J. Approx. Theory* **40,** 101–110.

MICHAEL, E. (1956), Continuous selections, I, *Ann. of Math.* **63,** 361–382.

PACKEL, E. W. (1986), Linear problems (with extended range) have linear optimal algorithms, *Aequationes Math.* **31,** 18–25.

PACKEL, E. W. AND WOŹNIAKOWSKI, H. (1987), Recent developments in information-based complexity, *Bull. Amer. Math. Soc.* **17,** 9–36.

PARTHASARATHY, K. R. (1967), "Probability Measures on Metric Spaces," Academic Press, New York.

PIETSCH, A. (1974), s-Numbers of operators in Banach spaces, *Stud. Math* **51,** 201–223.

PIETSCH, A. (1978), "Operator Ideals," Deutscher Verlag, Wiss., Berlin.

PIETSCH, A. (1987), "Eigenvalues and s-Numbers," Geest & Portig, Leipzig.

PINKUS, A. (1985), "n-Widths in Approximation Theory," Springer-Verlag, Berlin/Heidelberg/New York.

SMOLYAK, S. A. (1965), On optimal restoration of functions and functionals of them, *Diss. Moscow State Univ.*

STESIN, M. I. (1975), Alexandrov diameters of finite dimensional sets and classes of smooth functions, *Dokl. Akad. Nauk SSSR* **220,** 1278–1281.

TICHOMIROV, V. M. (1976), "Some Problems on Approximation Theory," Moscow State Univ. Pub., Moscow.

TRAUB, J. F., WASILKOWSKI, G. W., AND WOŹNIAKOWSKI, H. (1988), "Information-Based Complexity," Academic Press, New York.

TRAUB, J. F., WASILKOWSKI, G. W., AND WOŹNIAKOWSKI, H. (1984), Average case optimality for linear problems, *Th. Comput. Sci.* **29,** 1–25.

TRAUB, J. F., AND WOŹNIAKOWSKI, H. (1980), "A General Theory of Optimal Algorithms," Academic Press, New York.

WASILKOWSKI, G. W., AND WOŹNIAKOWSKI, H. (1984), Can adaption help on the average?, *Numer. Math.* **44,** 169–190.

WOŹNIAKOWSKI, H. (1986), Information-based complexity, *Ann. Rev. Comput. Sci.* **1,** 319–380.